D4.3 Interim Report on Reactive Behaviour and Haptic Techniques

Grant Agreement nr	856879
Project acronym	PRESENT
Project start date (duration)	January 1st 2018 (36 months)
Document due:	30/11/2020
Actual delivery date	30/11/2020
Leader	Inria
Reply to	Julien Pettre - julien.pettre@inria.fr
Document status	Submission Version

Project funded by H2020 from the European Commission

Project ref. no.	856879
Project acronym	PRESENT
Project full title	**P**hotoreal **RE**altime **S**entient **ENT**ity
Document name	Interim Report on Reactive Behaviour and Haptic Technique
Security (distribution level)	Public
Contractual date of delivery	30/11/2020
Actual date of delivery	30/11/2020
Deliverable name	D4.3 - Interim Report on Reactive Behaviour and Haptic Technique
Type	Report
Status & version	Submission Version
Number of pages	33
WP / Task responsible	Inria
Other contributors	-
Author(s)	Julien Pettre, Alberto Jovane
EC Project Officer	Ms. Adelina Cornelia DINU - Adelina-Cornelia.DINU@ec.europa.eu
Abstract	This deliverable reports our achievements along the following main directions: (i) provide designers friendly tool to extend in large amounts the believability, expressivity and richness of collective behaviors for agents; (ii) create new animation techniques to extend the expressivity of characters animation with possibility to perform on-line adjustments of the conveyed information through body gestures; (iii) explore new modalities for non-verbal communication with users with haptic techniques. The document also describes the explored techniques in their current version after slightly more than one year of developments within the project.
Keywords	Haptic Techniques and Reactive Characters, 1-to-n interaction scenarios, 1-to-1 interaction scenarios
Sent to peer reviewer	Yes
Peer review completed	Yes
Circulated to partners	No
Read by partners	No
Mgt. Board approval	No

Document History

Version and date	Reason for Change
1.0 24-10-2020	Document created by Julien Pettré, Inria
1.1 15-11-2020	Version for internal review (14 days before submission date)
1.2 30-11-2020	Revisions in response to review: final version submitted to Commission

Table of Contents

1 EXECUTIVE SUMMARY

Involved in interactions with users, we want the PRESENT agents to be capable of reacting in natural ways to users motions and behaviors. Reaction capabilities involve for example being attentive to users, show emotional reactions to their movements, adjust their behaviors to the changes in user position. It covers both cases of interaction where a user is interacting with a single agent (1-to-1 interaction), or a group of agents (1-to-n interactions). As the reaction capabilities of agents can be interpreted in many different meanings, PRESENT is focusing efforts on some main aspects : (i) provide designers friendly tool to extend in large amounts the believability, expressivity and richness of collective behaviors for agents; (ii) create new animation techniques to extend the expressivity of characters animation with possibility to perform on-line adjustments of the conveyed information through body gestures; (iii) explore new modalities for non-verbal communication with users with haptic techniques.

This deliverable reports our achievements along these three main directions, and describes the explored techniques in their current version after slightly more than one year of developments within the project.

2 BACKGROUND

This deliverable reports first advances made in WP4 to develop reaction capabilities of agents, as well as usage of haptic techniques to communicate through other sensory channels with users. These aspects are mainly explored by the Inria partner. We also report on studies that we have performed to better fix the requirements of both animation and haptic techniques.

3 INTRODUCTION

3.1 Main objectives and goals

When involved in an interaction with a virtual agent, any action of the users out of a set of predefined actions will demonstrate the lack of reaction capabilities of agents, starting from the most basic ones, such as when trying to startle them.

For this reason, PRESENT is exploring solutions to improve agents reaction capabilities in three two man ways:

1 In cases where users interact with a **group** of agents, by showing agents collective behaviors in complex scenarios, and how these collective behaviors adjust to users actions.
2 In cases where users interact with a **single** agent, by showing how the agents motion is adjusted to the changes in positions of the user, or scenario-guided indications to express, through body movements, specific intentions.

Since these two goals are aligned with the more general goal of improving non-verbal communication capabilities for agents, we finally add to:

3 Explore touch as a sensory communication channel for agents to convey information to users.

3.2 Methodology

The methodology applied to achieve these 3 goals will structure the report at hand.

Section 4.1 first describes the immersive setup we have designed, as well as the software architecture we have defined to satisfy our goals.

Section 4.2 describe our techniques to improve collective reaction capabilities for PRESENT agents. In this part of the work, we have identified the difficulty in designing new and potentially collective behaviors as a major bottleneck. To overcome these difficulties, we are developing an animation system based on the concept of *Interaction Fields* to elaborate new type of local interactions between agents, as well as with users.

Section 4.3 starts exploring expressive animation techniques through a perceptual study, where we evaluate the effect of non-verbal behaviour of virtual characters on people's response in the virtual reality environment.

Section 4.4 describes our technical approach to the problem of adjusting agents' animations to make them more expressive in reaction to users' actions or given scenarios. We elaborate an "expressive animation filter" that can process motion capture data, considered to be recorded in neutral conditions, and taint it with desired expressions.

Finally, Section 4.5 explores the importance of haptic rendering of contacts between users and virtual agents, and evaluates the usage of vibrotactile wearable devices, that meet the requirement of immersion with mobility and light equipment.

4 MAIN CONTENTS OF THE DELIVERABLE

This section is made of five main parts. The first one, section 4.1, provides an overview the PRESENT setup for enabling interactions between users and virtual agents, including through touch by using haptic techniques, and through which PRESENT will develop and evaluate agents' reaction capabilities.

4.1 Haptic Techniques and Reactive Characters

4.1.1 Overview of the setup and architecture for reactive characters.

The PRESENT project is primarily concerned with the immersive experience of a user among his digital alter-egos. There are many technologies to achieve this immersion, and to enable interaction between users and virtual agents. The two section below describe the immersive technologies we have selected for immersion, as well as some specific elements of the software architecture to achieve interaction with agents.

4.1.2 Immersion technologies

This section describes the employed immersion technologies to enable users and virtual agents sharing interactive experience. In view of the addressed topics, immersion techniques have been selected to allow:

- Immersion with multimodal sensory feedback combining a minimum of visual and haptic feedback.
- Capture user-generated motion and sound to study the reactive capabilities of virtual humans.
- Freedom of movement and displacement of the user, to allow in particular the study of 1-to-n scenarios based, as described below, on the dynamics of global positions in the environment.

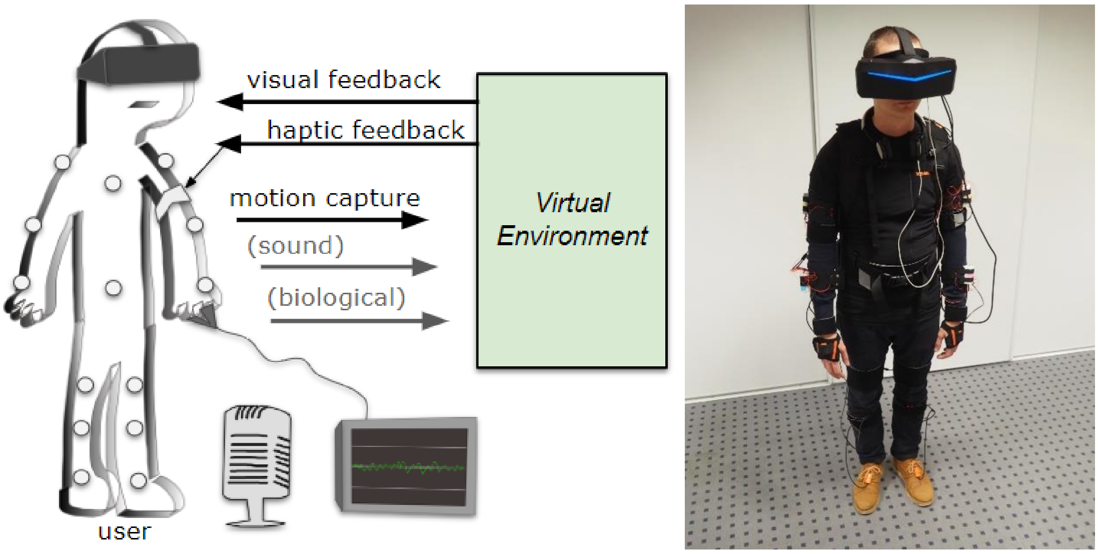


Figure 1 – Illustration of the immersion modalities and technologies for PRESENT

For the needs of PRESENT's work, Inria has therefore gathered and tested a set of immersive technologies composed of the following elements (cf. Figure 1):

- For visual feedback, the use of HMD allowing immersion without strong spatial limitations to move. Here, the PIMAX model offering a large visual field was chosen.
- For haptic feedback, the use of custom unobtrusive wearable peripherals providing vibrotactile feedback.
- For motion capture, we have two systems used in combination: a system using inertial sensors (XSens suits) and an opto-electronic system.
- The capture of sound is done by a simple microphone, the intention being to be able to detect sound peaks only, to allow a reaction of virtual humans. Note that in the current version of immersive setup, this modality is not activated nor used in practice.

4.1.3 Architecture for reactive characters

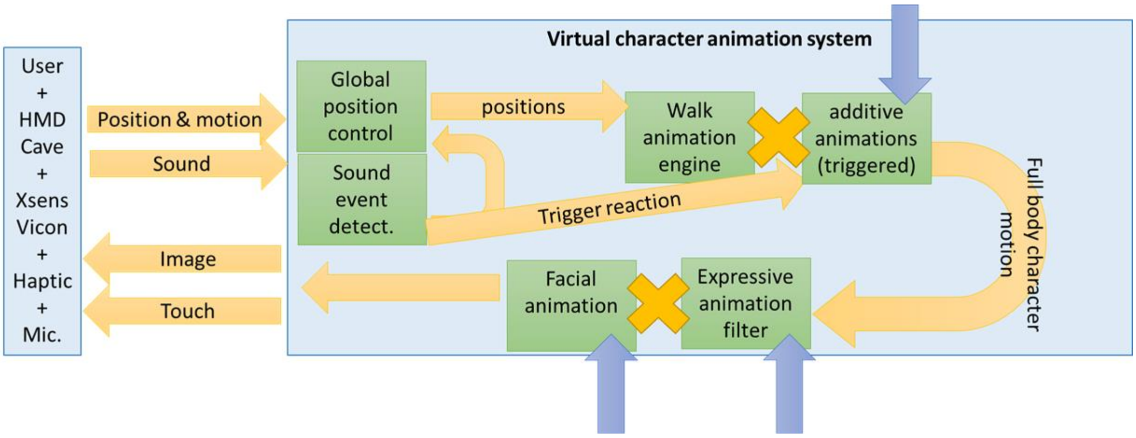


Figure 2 – Architecture for reactive characters

The PRESENT project develops the reactive capabilities of virtual agents with two categories of target scenarios: interaction between a user and a single agent, for which we want to have detailed body reactions to users, as well as interactions between users and groups of agents, where for example users can get surrounded by agents. Could they be involved in a 1-to-1 or 1-to-n interactions, we focus on reactive capabilities for agents involved in non-verbal communication modalities, making links with other WP4 tasks, and more precisely through body motion and gestures. The architecture main blocks are:

- The user sensing system (left block): user is motion-tracked so as to enable agents' reaction to their gestures and changes in position. Sound is also recorded to make agents capable of reacting to sudden sound. However, in the current version of the developments, sound modality is not yet considered. Position and motion information is constantly broadcasted to the virtual agent animation system.
- Two components consider respectively the sound and motion data flows. As explained above, today, sound is not actually processed. Our plans is to perform later a sound event detection system that would trigger reaction for agents. The user's motion information is however processed early in the system by the "global position control" component. This component is in charge in a 1-to-n interaction scenario to control all the virtual agents' positions, so as for example, to go towards the user or at the opposite move away from it, avoid it, hide from him etc. The 1-to-n interaction techniques are detailed in Section 4.2.
- Since only the global positions of agents are controlled, agents' positions are sent to the "walk animation engine system". This component uses state-of-the-art animation techniques to turn agents global trajectories into complete whole-body animation (performing walk motions along global trajectories).
- At the stage of performing walking animation for virtual agents, additive animations can be triggered by an external animation control system. Such connections to the virtual agents animation system enables integration in the PRESENT architecture.
- The resulting animation is then sent to the "Expressive Animation Filter" component. This component is an essential part of PRESENT contributions. Indeed, this filter's role is to further edit the resulting animation so as to comply to 1-to-1 interactions with the user. Simple example is to edit the gaze direction so as to look at the user, or to adjust body or head orientations toward the user. More interestingly, as described in section 4.4, agents' animation can be adjusted according to a desired scenario and control the expressivity of motions.
- Finally, still to fit in the PRESENT architecture, facial animation (out of the current deliverable activities) can be added to the character.
- Two sensorial feedbacks are delivered to the user, respectively through the visual and tactile sensory channels.

4.2 Agents' Global Position Control: 1-to-n interaction scenarios

In this section, we describe our system to control agents position to react to users presence in the virtual environment, as well as his behaviours. Our choice is to control these positions following a paradigm in crowd simulation techniques, where the behaviour of each crowd agent is controlled through models of local interactions, that dictate how one's motion is influencing others neighbour's motion. However, existing techniques do not allow for rich interactions scenarios. Our effort in PRESENT is thus dedicated to create an animation system with designers' friendly techniques to achieve complex models of interactions in an intuitive way.

Many applications of computer graphics, such as cinema, video games, virtual reality, training scenarios, therapy, or rehabilitation, involve the design of situations where several virtual humans are engaged. In applications where a user is immersed in the virtual environment, the (collective) behavior of these virtual humans must be realistic to improve the user's sense of presence. As part of realism, expressive behavior appears to be a crucial aspect. For example, Slater et al. (2006) showed that the expressive behavior of an audience in VR had a direct impact on the speaker's performances and perception of themselves. This work concerns the motion through an environment of virtual humans. In the area of crowd simulation, collective behaviors are typically simulated using models based on forces (Helbing 1995), potential fields (Treuille 2006) velocity selection (van der Berg 2011), or vision (Lopez 2019). However, those techniques lack expressiveness and do not allow to capture more subtle scenarios (e.g., a group of agents hiding from the user or blocking his/her way), which require the ability to simulate complex interactions. As subtle and adaptable collective behaviors are not easily modeled, there is therefore a need for more intuitive ways to design such complex scenarios.

4.2.1 Context

The category of 1-to-n non-verbal communication scenario considers the interaction between a user and several virtual humans. We are interested in the case where virtual humans are part of the same group and interact with a user, who may feel as part of the group or not, and whose decisions and reactions may be influenced by the group. Several theoretical works regarding social interactions have motivated the design of such scenarios.

According to Wilder (1986), "persons organize their social environment by categorizing themselves and others into groups". Three categories have then be described (1) there is no relation between the perceiver and the group, (2) the perceiver is a member of the group (in group), and (3) the perceiver does not belong to the group and compares with his/her own group (in group/out group). This categorization implies notions such as similarity, homogeneity and differences and would allow individuals to simplify their social environment and predict future social behavior. In addition, previous works on conversational group (Kendon 1990) have shown that the relative position of the members of the group, which can be considered as an in group situation, are set in a way that each member of the group has a similar shared space with direct and exclusive access. Kendon refers to the F-formation system which describes this "spatial-orientational behavior" and that can be dynamically adapted so as to include another person in the group. Recently, Cafaro et al. (2016) have used closely related concepts to design believable virtual agents in small conversational groups (static condition) exhibiting nonverbal behavior. Authors manipulated the relative position of each individual (group formation) as well as interpersonal attitudes (friendly vs. unfriendly). They defined the "in-group attitude" which was directed towards the member of the group and observable by the user as not being a member of the group, and the "out-group" attitude where an overall attitude was expressed towards the users' avatar approaching the group. Results showed that out-group attitude has a main impact on social presence and that proxemics, i.e., interpersonal distance, was affected by the in-group, out-group attitude.

In line with these studies, we would like to extend the design of interactive and expressive virtual humans to dynamic situations. Especially, we aim at modulating the non-verbal expressivity conveyed by virtual humans through their collective motion. Collective motion emerges when individuals achieving the same goal interact with each other. In this context, simulating an expressive and collective motion consists in defining the respective position of each virtual human in time so as to convey a certain amount of unity between the members of the group. The modulation of the group expressivity will be achieved through their motion and final configuration relative to the one of the user. Based on this in and out group concepts as well as group formation principles, we designed in collaboration with CREW, a surrounding scenario that will allow us to manipulate the valence of the situation from the user point of view.

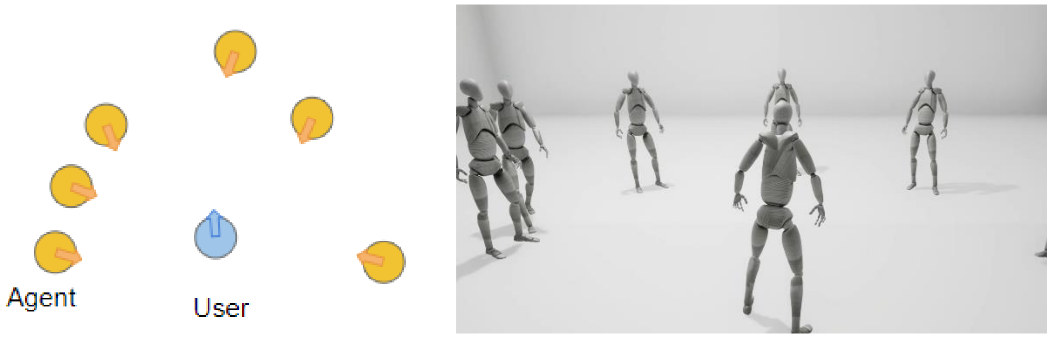


Figure 3 - Illustration of 1-to-n scenarios, where a user is surrounded by a group of virtual agents

Example of such scenario is illustrated in Figure 3 (co-designed with CREW): a user is approaching a group of virtual agents. We elaborate on the reactions and the expressions this group of agents could exhibit. In terms of reactions a large palette of body motions may convey to the user the fact that its presence among virtual humans has triggered events. This could be about making eye-contact, turn bodies toward the user, move toward the user, move away from the user, leave him room to join, etc. Each can be categorized as being neutral, positive or negative (valence). The intensity of such a reaction is adaptable. Synchrony and propagation in reactions will convey the reaction collectiveness.

4.2.2 Overview of the PRESENT system for local interaction design and simulation

We here give an overview of the "Global position control", which comprehend 2 parts: 1 technique to design new kind of interactions in an intuitive way, and 1 part to simulate them. As explained above, local interactions between agents is the key component of our approach to model and simulate emergent collective behaviors. Many methods exist allowing us to apply a navigation algorithm to a group of agents by modeling their behaviors and interactions: most local interactions are expressed as an energy minimization problem, which requires to express local interactions through energy terms. It turns out that the need of formulating such energy functions can be a bottleneck in the development of models of more complex interactions, where no simple function can transcode human behaviors. For this reason, PRESENT is working on intuitive and expressive methods to describe interactions. This is the main goal set for our approach called "Interaction Fields": to enable users to create interaction behaviors between agents in an intuitive and generic way to fit the needs of more complex scenarios.

Examples of more complex interactions we are exploring are: remain in view of someone, approach someone to stop him and talk to him, or the opposite stay out of the field of view of the immersed user and hide behind obstacles to do so. They could as well always try to stay behind one another's back to hide. On the contrary, the goal could be to jump in front of the user to block the passage as soon as the user wants to go in a certain direction.

To model those precise local interactions in an intuitive way, we are exploring the possibility of directly drawing the interaction and thus, visually defining it. By drawing a vector field centered on its source (an agent, an object, an obstacle…), those vectors can be applied to the other agents depending on their position in the scene. According to the relative position between two agents, the influence they have on one another is given by their vector fields.

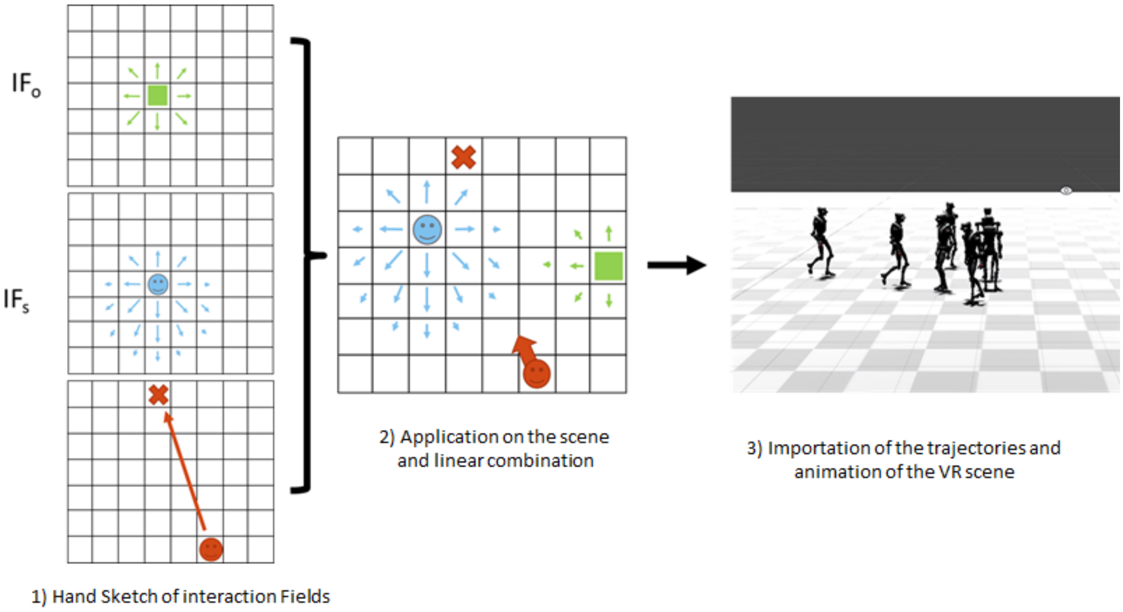


Figure 4 - Main steps of the Interaction Fields method

The method is based on three stages illustrated in the Figure 4. The first stage is the drawing of the Interaction Fields in a discretized space and centered on their sources. A few control vectors can be drawn and the fields are extracted using interpolation techniques. In Figure 4(1) we consider a vector field from an agent (IFs) and one from an obstacle (IFo). We considered those vector fields affecting one agent (in red) whose goal is to reach the red cross. Then we combine linearly the fields and apply them in a scene. The vectors applied to the red agent is a sum of the vector fields of the blue agent (IFs) and the green obstacle (IFs), see Figure 4(2). From the vector of each cell x, we can get the acceleration a of the agent. With the acceleration of each agent it is possible to get the trajectories on the scene online and to animate them in Virtual Reality, body animation like walk or gesture can be added on top of the trajectories see Figure 4(3).

4.2.3 Designing Interaction Fields

In line with crowd-simulation research, we simulate the environment as a 2D plane, and agents as disks whose positions and velocities are 2D vectors. The task of an Interaction Field (IF) is to describe how agents should move through the environment in the presence of another agent (which we call the source agent). An IF is a square with the position of the source agent at its center. Each point in this square prescribes a velocity that other agents should use when they are located there. In practice, we store an IF as a grid, and we compute velocities via interpolation between grid cells. During the simulation, if an agent A is near a source agent S, it will translate and rotate the IF to match the current position and orientation of S. Agent A will then choose the velocity that the IF prescribes for A's current position.

4.2.3.1 Control vectors to design Interaction Fields

We developed a simple interface enabling a user to intuitively draw the IF for each agent. In this interface, the user can spawn a new agent and its corresponding control vectors are intuitively drawn directly around the source agent. To do that there are three stages:

1. Positioning of the source of the Interaction Fields. The pole of interaction (agent on the left or obstacle on the right) can be deleted or displaced.
2. Adding the control vectors drawn directly by hand, those control vectors can be curves or lines. The control points of the curves (shown in yellow) or lines (start point and end point of the line) can be moved. The lines or curves can be entirely displaced. In the case of the curve, the user can increase or reduce the number of control points to get more precision on his drawing. In blue, you see the hand drawn line created directly by the

mouse movement. The number of lines or curves is unlimited. When a control point of a curve is displaced, the curve is updated using a Bézier curve. When the curve is set, the control vectors used to extract the Interaction Field are:

a. In case of the line: simply the vector of the line
b. In the case of the curve, the curve is divided in equal segments (defined by the control points given by the user), that each represent a control vector following the curve direction. The amplitude of the control vectors is decided by the width of the curve that can be modified by the user as well.
c. The last stage is the interpolation of the control vector to create a vector field. To extract the vector field from the control vectors given by the user, we use interpolation. This part will be thoroughly described in part 2) below.

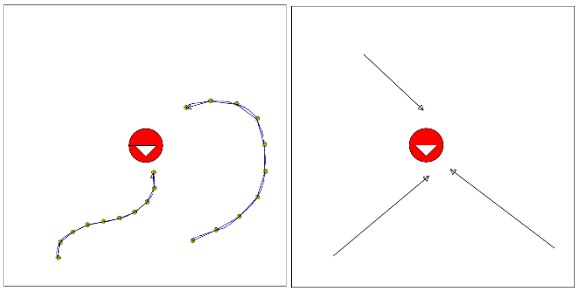


Figure 5 – Representation of an interaction field being designed with, in red, the source (an agent here) and some control vectors drawn by the designer to orient the interaction field accordingly.

4.2.3.2 Computing Interaction Fields

We use an inverse distance weighting interpolation technique to compute interaction fields from the control vectors. To this end interaction fields is first decomposed in cells. The number of columns and rows can be chosen by the user which will have an impact on the size of the final field. Each of the control vectors obtained from the user's drawing are defined by a start and an end point. They then will be appointed to the cell that contains the start point. Be:

- $(u_i)_{i\in[1,N]}$ the ensemble of control vectors obtain in cell x_i.
- $u(x)$ is the vector interpolated inside cell x.
- $u(x) = \frac{\sum_{i=1}^{N} w_i(x)\, u_i}{\sum_{i=1}^{N} w_i(x)} \; if \; d(x, x_i) \neq 0 \; \forall \, i \in [1, N]$
- $u(x) = u_{i,} \quad if \; \exists i \in [1, N] \,|\, d(x, x_i) = 0$

Where $w_i(x) = \frac{1}{d(x,x_i)^p}$ and $p \in \Re^* |\, p \leq 2$ the power parameter.

$$d(A,B) = |x_A - x_B| + |y_A - y_B| \; is \; the \; manhattan \; length \; between \; the \; point \; A(x_A, y_A) \; and \; B(x_B, y_B).$$

The power parameter p allows to decide how much the distance should impact the interpolation, the more p is large, the more the cell is going to be impacted by far way control vectors. After several example, we noticed the importance of this parameter, and in order that the resulting field would be impacted by the overall control vectors but more so by the closest ones, a value of p superior to 1.5 is preferred.

4.2.3.3 Examples

Currently, a version of the Graphical Interface to draw the Interaction Fields has been developed. The key-idea of this interface is to let users draw some guidelines according to the position of the source of the field. Vectors of the field are directly oriented according to those guidelines, whilst the whole Interaction Field is then resulting from interpolation. Some examples are visible in Figure 6. The three top figures show an agent (yellow circle) with which other agents are going to interact, the input curves drawn by the user (blue), and their resulting control vectors (red) in three representative trials. The three bottom figures show the resulting IF for the same three trials after interpolation. Here the source is an agent, represented by a yellow circle.

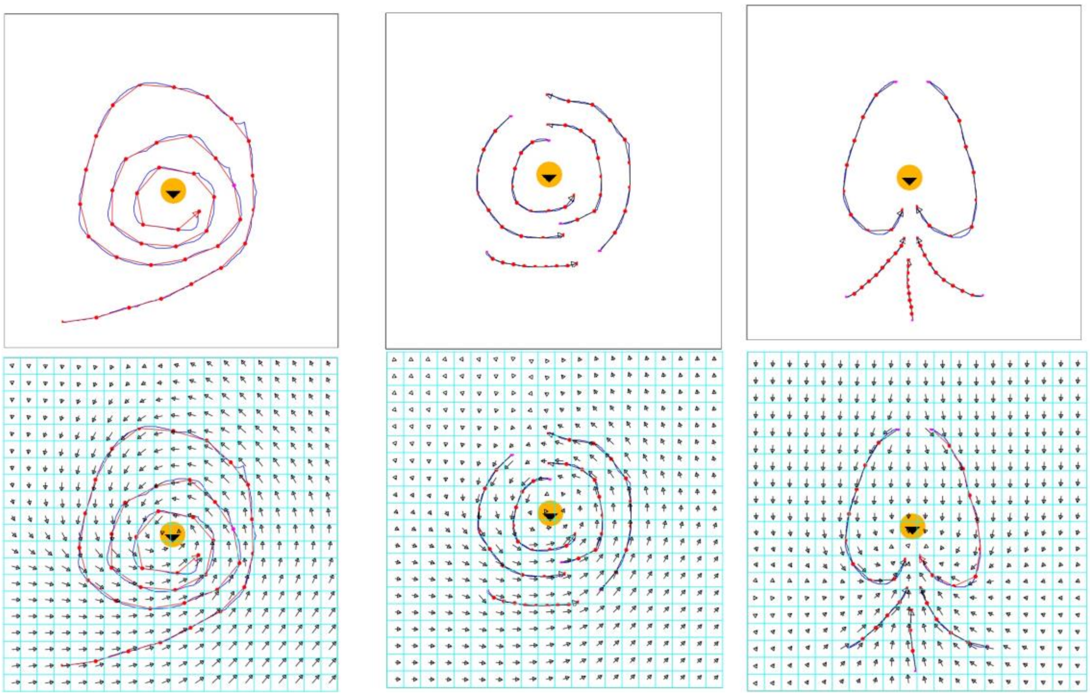


Figure 6 – Examples of Interaction Fields

The first example of Figure 6 shows a possible Interaction Field design for a spiral shape trajectory. The agent influenced by the field, would start at the border if the field and turn around the source of the field until the source of the field. The second image shows the example of a circle shape Interaction Field where the influenced agent would simply turn around the source of the field. For this case, one could imagine drawing several control vector curves, as shown, or just a continuous circle control vector around the source, similar to the first example. For the last example, the idea is to make the influenced agents, neighbors of the source agent, go in front of the source agent. For this, several control vector curves are necessary and the direction of the source agent is important.

4.2.4 Agents motion control with Interaction Fields

The IF then impacts all the surrounding neighbor agents of the source to define their interaction with the source of the IF. Then, using interpolation techniques, the Interaction Field is deduced from the control vectors. Each agent can be a source of her/his own IF. The fields are then combined in a simulation and applied depending on the relative position of the agents to the sources of the fields.

4.2.4.1 Applying the fields

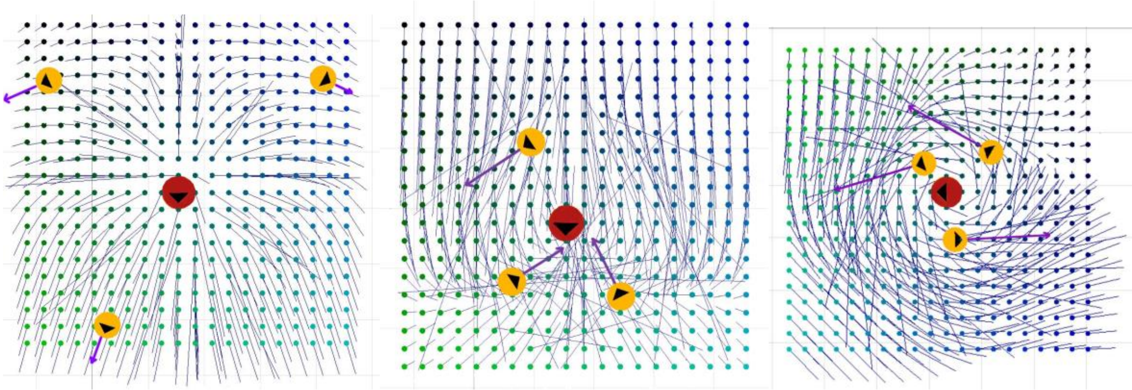


Figure 7 - Examples of IF applied by the red agent during a simulation: each yellow agent is moved by the combination of their neighbours' IF (resulting vector in purple). Left) repulsion IF applied by every yellow agent in the pictures. Middle) circle-shape IF designed in Figure 6, middle. Right) the IF applied during simulation designed in Figure 6, right.}

An Interaction Field (IF) is applied to a given agent according the scene setting (in Figure 7, the red agent). Several Ifs can apply to one single source. Several fields can hence impact one agent. Combining the Interaction Fields is an important point that must be addressed with care. In each frame of the simulation, each agent identifies its neighboring agents, and then each agent computes a new velocity according to the interaction fields that it applies to each neighbor.

- For each space cell x_k of each Interaction Field IF_k the vector of the interaction k is:

$$\vec{x}_k = IF_k(x)$$

- Then we combine linearly the fields and apply them in a scene. The final vector of cell x of all Interaction Fields applied to one agent *i* is:

$$\vec{x} = \sum_k \overrightarrow{x_k} = \sum_k IF_k(x)$$

- To enable more subtlety on the combination of the fields, we define weight w_k for each IF_k. We have:

$$\vec{x} = \sum_k \overrightarrow{w_k x_k} = \sum_k w_k IF_k(x)$$

- Then, to find the acceleration of agent i of mass m impacted by the sum of the IFs at each frame of the simulation, we define

$$\vec{x} = m_i \overrightarrow{a_i}$$

The orientation and the position of the Interaction Field is updated according to its source agent. Figure 7 shows the resulting simulations of two of the previous designed Interaction Fields (Figure 7(middle) and Figure 7(right)). The figures are screenshots taken during the simulation involving four agents moving in the same scene, where the three yellow agents' IF impact the surrounding neighbors by a simple repulsion IF (see Figure 7 (left)) representing social distance. The yellow agents' IFs are not displayed for more clarity. The agent in red is moved using the keyboard during the simulation and applies the visible IF to the three other agents. The resulting motion vector is shown in purple for the yellow agents.

4.2.4.2 Interaction Fields types and their combination

At this stage of this work there is three types of IF:

1. Velocity Interaction Fields: impact the velocity of the agent undergoing the field at each frame.
2. Orientation Interaction Fields: impact the orientation of the agents inside the field. To control the orientation and the positioning of the agent allows for a more detailed and realistic scenario. The field is designed exactly the same way the velocity Interaction Field is designed.
3. Parameter-dependent Interaction Fields: interactions may depend in some parameter value. For example, two fields must be designed beforehand in our interface. One field is designed for a static source, and the other for a moving source at full speed. To apply those fields together, we create a new temporary field, IF_{result} based on an interpolation of the two fields. For each space cell x_k of each Interaction Field IF_{static} and $IF_{dynamic}$ the vector of the interaction k is in field IF_{result} :

$$IF_{k\ result}(x) = \frac{IF_{k\ dynamic}(x).|\vec{v_S}| + IF_{k\ static}(x)(v_{max} - |\vec{v_S}|)}{v_{max}}$$

   Where $\vec{v_s}$ is the current velocity of the source agent and v_{max} is the he/her maximum speed.

Having a various set of different IF types allows us to address a more important number of scenarios. The Interaction Fields and the resulting behaviors are then more controllable and more intuitive to design. As a result, our current set of IF types will presumably increase even more, to fulfill the need to keep a generic framework.

4.2.4.3 Example

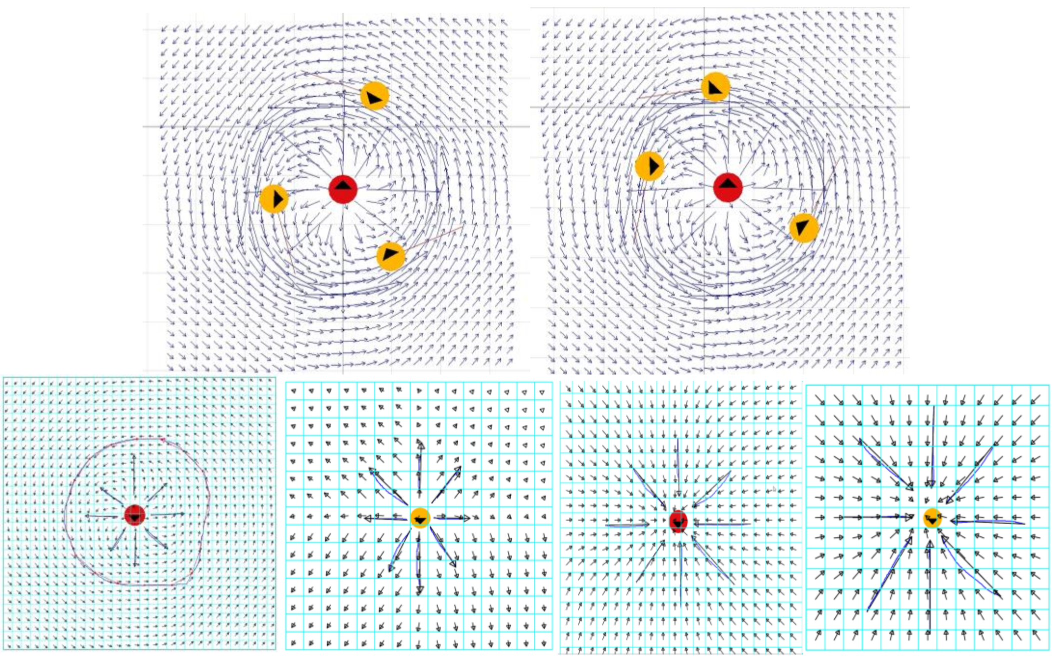


Figure 8 – (top) Example of 3 agent interacting a user (red agent). (bottom) This interaction results from 4 Interaction Fields: circling around the red agent, repulsing one another, orienting toward the red agent, and getting attracted toward the red agent.

Figure 8 illustrate an example where a complex case of interactions, where 3 agents circle around a red agent and remain oriented towards it. This interaction results from 4 Interaction Fields, each being intuitive to design.

4.2.4.4 Animation

Once trajectories of agents interacting with each other have been generated thanks to the Interaction Field method, it is however still necessary to animate their body motions. To this end, we use a plugin using the technique of Motion Matching. Motion Matching allows, given a defined library carefully chosen, to select the animation from this library and to achieve a transition between them. To choose the right animation, the plugin uses predefined key points of the agent skeleton that follow the trajectories (like the ankles).
In order that the functionality of the developed interactions could be used by other project partners, progress in connecting the hand sketched interaction fields with Unreal have been made. Collaboration with the CREW partners who will be able to implement the Interaction Field approach to animations for their interactive scenarios, are undergoing. The OSC plugin of Unreal was used to broadcast data between the interaction field and Unreal. OSC or Open Sound Control is primarily used to network generic audio data between clients, but in our case, it was used to send position data from our software to Unreal.

4.3 Expressive animation filter: 1-to-1 interaction scenarios

4.3.1 Context

The 1-to-1 non-verbal communication scenario category considers interactions in dyadic situations, i.e., between a user and a virtual human, where interaction cues are conveyed through voluntary and involuntary body gestures and postures as well as eye movements. In this category of scenarios, we focus on the animation of expressive and reactive body motion, that are fundamental, not only for non verbal local interaction applications, but also to serve verbal communication and collective response (cf. section 1-to-N non verbal communication). Our goal in these scenarios is to demonstrate characters that are able to individually react to the user to efficiently communicate with them, e.g., using body gestures and eye contact to attract the user's attention in a natural manner. Our approach relies on the main features that characterize non-verbal interaction behaviors (Juslin et al., 2005), namely 1) Proxemics, 2) Head, body and limbs posture and motion and 3) Gaze. During this first period, we focused on proxemics and body movements animation.

4.3.1.1 Proxemics

Proxemics is the study of people's perception and use of space (Hall 1963). The exploration of this field of study emerged in the 1960's as an interdisciplinary approach to understanding complex human behavior in crowds. The mechanisms which respond to density are physiological, indicating higher stress levels when there is overcrowding (Christian 1961), they are cultural, since no universal patterns of the use of space was found across cultures (Hall 1963), they are dependent on the subjects - an observer's and confederate's gender (Brady 1971), behavior (Argyle 1969), attractiveness (Kmiecik 1979) and other contextual factors. In this context, Remland et al. (1995) have shown that interpersonal distance, orientation, and touching are of interest to capture information regarding the spatial separation between interactants.

4.3.1.2 Virtual Reality

From a methodological point of view, it is worth pointing out that virtual reality provides a unique and valuable tool for proxemics researchers. Observational methods in physical reality allow for little or no experimental control, confederates who may behave inconsistently, and employ projective measurement techniques (Bailenson 2003). Virtual reality allows investigators to maintain complete control over virtual human representations' appearance, behaviour, and environment while ensuring a high degree of ecological validity. We then conducted two experiments to study proxemics in virtual reality in the frame of social interactions. The first one, which is directly related to expressive body movement animation applications, focuses on the perception of gender in motion in a proximity experiment, (see Section Proximity). The second deals with the evaluation of haptic rendering of contact when a user interacts with virtual humans (see Section Influence of haptic rendering of contacts on user behaviours).

4.3.1.3 Body motion:

Regarding posture and body motions, we first provided effort in the technical development of expressive animation pipeline. We aim to go beyond the limitations of current animation techniques, which are based on data-driven motion synthesis (which means smartly pick the right animation at the right time [motion graph, motion matching]), those rely on pre-recorded motions.
Our approach wants to coherently combine in a single pipeline motion synthesis methods, described before, with a procedural motion editing approach. We are currently working on the design of filters that will modify a given input motion generating an expressive output motion. Our approach relies on the visual perception of the user, where the spatial temporal control of the animation is performed in relation with the user's field of view (see section *Expressive animation filters*).

4.3.2 Proximity study

As part of the project WP4, a preliminary investigation into the effect of non-verbal behavior of virtual characters on people's response in the virtual reality environment was conducted. We focused on how biological motion, captured from an actor and applied to the virtual character, influences people's response to this character. When designing high fidelity agents in the scope of the PRESENT project, understanding how the information from motion influences observers' response, is important.

We investigated two components of movement which could potentially affect people's response to virtual crowds – the gender and attractiveness of motion. This was hypothesized based on previous studies which found that both gender (Iachini, 2014) and attractiveness (Kmiecik, 1979) impact human interaction, but it was never established if motion (which carries both gender and attractiveness cues) could have an impact on its own. We measured the response by using proximity, a measure, which is associated with the previously mentioned study of personal space (proxemics). We created a virtual environment (see Figure 9) with a wooden mannequin model, to which walking motions of 20 actors (10 male and 10 female) were applied. Participants were embodied in the same virtual model, watched all motions in randomized order and were asked to stop the approaching character as soon as they felt uncomfortable with their proximity. The distance between the character and the participant at the time of the stopping was recorded. The participants were also asked to report how attractive the motion of the character was and recognize the gender of the actor from the observed motion.

Figure 9 - The view of the participant from the experiment - a 'genderless' model walking towards the participant, embodied in the environment.

The results of this study show that biological motion, when applied to virtual characters, has an effect on the viewer's response. Particularly, we found:

- people feel more comfortable with an approaching character, whose motion is perceived as 'attractive';
- female observers report larger proximity to the characters than male observers. This shows that the gender of the observer has an impact on the use of personal space with virtual characters;
- we found no effect of the virtual character's gender on the proximity, even though it was expected.

The results of our study provide a contribution to the understanding of interactive behavior in virtual environments between observers and virtual characters, as well as a useful guideline for designing and manipulating motion of virtual humans in the scope of the PRESENT project.

4.3.3 Expressive animation filters

As introduced before, the *Expressive animation filter* aims to modify a given animation to maximize the generated expressivity directed to an interlocutor, in the context of a 1-to-1 non-verbal interaction.

It is designed to be placed as a middleware inside a fully operating character animation pipeline (Figure 2, also detailed in Figure 10), right after a motion is picked to be displayed. The goal of these *expressive animation filters* is to work on the finer animation details to enable the virtual characters to be more expressive and therefore engage the user. More specifically, our approach explores the use of animation editing tools, such as motion warping (Witkin et al. 1995), to exaggerate or mitigate expressive features of the motions. By knowing the user's current position, we will be able to continuously evaluate the perception he has of the virtual character and adjust the motion accordingly. Our objective is to give a simple control over the filter, by setting the desired expressive result (for example: more/less aggressiveness, express dominance, etc…).

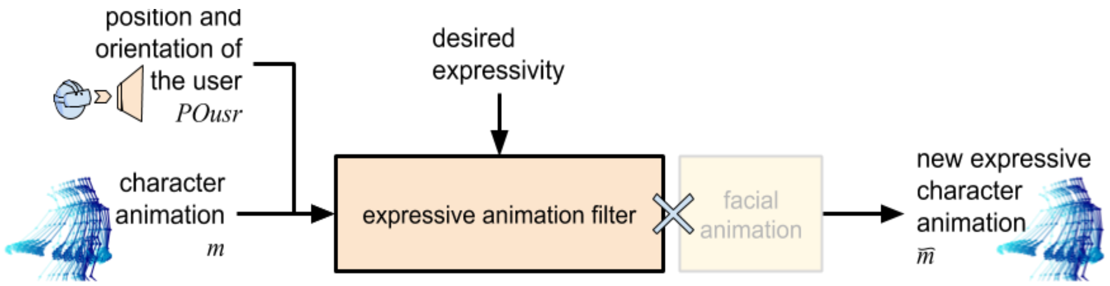


Figure 10 - The position of the *Expressive animation filters* inside the structure of a general character animation pipeline.

4.3.3.1 Structure

Here we introduce the algorithmic structure and the notation of the filter.
Our inputs are:

→ a full body character motion *m*, defined as an interpolation of keyframes over time $m = \{p(t)\}$ where $p(t)$ is a pose at a specific time t.

→ the observer position and orientation over time *POusr(t)*,

→ additional control's parameters (like the desired expressivity value).

← The output is a new warped motion $\underline{m} = \{\underline{p}(t)\}$.

In order to fulfill these requirements we have defined two classes of functions: (i) the motion features function MF, that numerically evaluate the motion from the observer perspective, and (ii) the warping units WU, that are in charge of modify and perturbate the original motion. These WUs can be further split on two different levels: • pose warping units, • time warping units both in charge of generating small variations of the motion, without affecting the original "essence" of the animation. The first one acts on positions of static poses, and the second one on time-related attributes (velocity, acceleration, etc…). Additionally pose warping units can be extended introducing more invasive modification intended as extreme cases of the local warping, in order to generate poses that cannot be achieved with simple perturbations of the original motion (for example moving the hands to cover the head in case of an incoming threat).

Combining these two classes of functions (motion features function and warping units) we structure the internal logic of the *expressive animation filter*, where we evaluate the current state of the motion with the MFs, and then we adjust it and modify it accordingly with the WUs. Different approaches are possible, here we will introduce one based on the definition of a Jacobian matrix, but we plan to explore and compare different methodologies.

4.3.3.2 Motion Features

As we said, motion features functions evaluate numerically the visual perception of a character animation. Each specified function is dedicated to a specific visual cue (positions, lengths, angles, areas, velocities, accelerations, etc…) of specific limb parts (head, arms, hands, etc..).
A motion feature function is defined as $MF(m, POusr)$. It takes as input a motion m and the position and orientation of the user $POusr$. Then the computation pipeline is straightforward:

1. get and aggregate the desired data from the position of a set of selected joints .
2. project this on the field of view of the user.
3. compute the feature on the 2D projected plane, accordingly with the user perception.

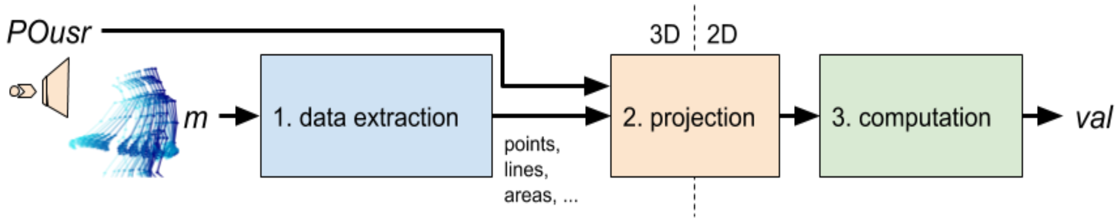


Figure 11 - General computation of the motion feature value. From the motion m firstly we extract the relevant joints, at a specific frame/s, and then we combine them together depending on the feature (distances, areas, velocities, ...). The second step is the projection in the user viewing plane, so depending on his position and orientation $POusr$. Finally the 2D data era elaborated to extract the final value.

The output of the motion features function is an adimensional scalar value val with no related meta-information. From previous works in psychology and animation (Randhavane et al. 2019) we know that there exists a link between visual cues in human movement and expressivity (or perceived emotional state). In order to fill this missing gap between visual cues and correspondent expressivity, we are working on formalizing a list of motion features (distances: positions and relative positions of limbs, angles between limbs, areas formed by different joints, amplitude of a movement, velocity, acceleration, jerkiness of a limb, ...) and validate their impact on expressivity through participants studies, similarly to the ones performed in the Proximity Study (Section 4.3).

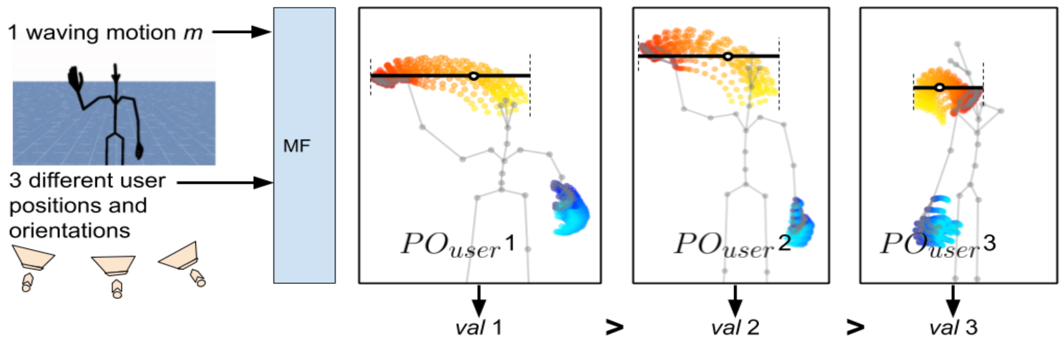


Figure 12 - Example of MF evaluating the amplitude of waving animation. The animation is the same in the 3 cases, what is changing is the point of view of the user, and so the resulting amplitude in the 3 cases, ordered from the bigger to the smaller.

4.3.3.3 Warping Units

The second family of functions is the warping units (WUs), as we saw before, these are a set of modifiers for the original motion, and their goal is to affect the visual perception of the motion, and so consequently generate observable variation in the values computed by the motion features functions (MFs). Each WU is localized and attached to a specific set of joints. The original idea is derived from the action units introduced for facial animation (Facial action coding system Ekman et al. 1978) defined as coordinated movement of a set of facial muscles. This concept has been previously applied to the body in the body action and posture coding system (BAP) (Dael et al. 2012), where body action units are described as a set of limbs motions relative to the axial structure of the body (sagittal, vertical or transverse) and with a detailed classification (emblems: action with a intended meaning, manipulators: gesture that manipulate other part of the body and illustrator: conversational gesture). We have abstracted our definition from the posture, and relate the warping to arbitrary references. Another important inspiration is the synergy, defined in neuroscience as coordinated muscle activation. Synergy is generally used in robotics, for replicating human gestures, like grabbing objects (Santello 2016), we define WUs as coordinate joint movements. Inside warping unit we identify different type of functions:

Pose warping units: a pose warping unit acts only on static pose and generates a combined rotation of a set of joints, it is defined as a triples (<joint, axis, weight>): (i) a joint (ii) an associated reference rotation axis, that can be set internally from the body or externally, and (iii) a weight. The action is then an ordered weighted composition of these rotational triples. From this general definition we close in to specific cases: spine-torso flexion relative to the hips orientation, shoulder opening/closing accordingly to the top spine, arm rising/declining relative to the ground (absolute reference system). The modifications are applied as perturbation around the input pose, selected as the starting state. In any case, in order to avoid undesirable results, like unrealistic posture, it is possible to add a layer of constraints on the joint's degree of freedom (dof), depending on the body structure.

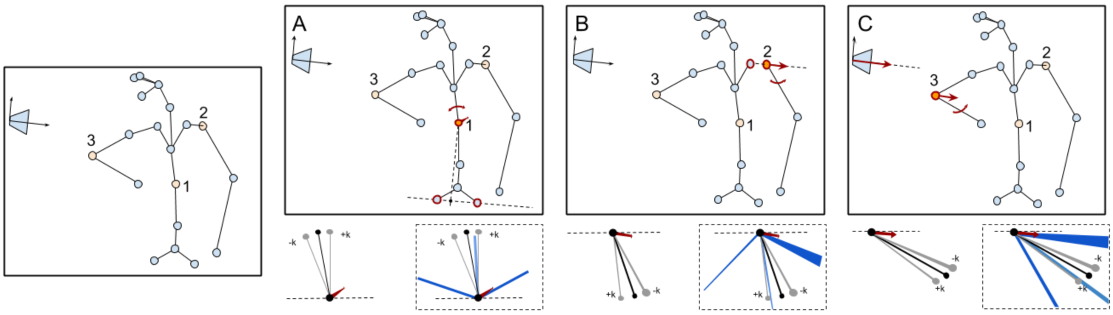


Figure 13 - Three examples of axes definition for a joint dof used in pose warping units: on the left we have the static pose with the 3 joints (1,2,3) highlighted. In (A) the axis of joint 1 is the normal vector of the plane, defined with 3 positions (in red), and through the joint 1. In (B) the axis is defined from two joints positions (in red). In (C) the axis is defined from an external reference, the user 's view direction. Under each case, on the left: there are the representations of the modification caused by the variation on the *k* value (*+k* or *-k)* and the weight, additionally on the right there is an example of how the modification can be restricted on a range (dark blue) and a zero case (light blue).

Time warping units: similarly to the previous one unit operates on a specific set of joints, but instead of modifying the rotation of a single pose, it applies modification on multiple poses simultaneously affecting time dependent features (like velocity and acceleration).

A general warping unit is simply defined as $WU(m,k) = \underline{m}$ getting a motion m and a k parameter that is the one tuning the intensity and the direction of the variation, $k = 0$correspond to the original motion. The result is the modified motion.

4.3.3.4 Expressive Filter Core:

The expressive filter core is a combination of motion features functions (MFs) and warping units (WUs), targeting the evaluation and control of visual cues in the human motion. We are trying to solve an optimization problem, where we have a desired target result t (desired motion features, coming from the selected emotional state) and a current state s (current motion features, computed with the MFs functions applied over the current motion), the result is a combination of WUs applied on the current motion to minimize the error $\varepsilon = t - s$ and so moving the current state to the desired one. The algorithm operates in a space defined by the motion features functions (and so dependent on the motion and the user field of view) and the movement around this space are parametrized by the warping using WUs.
We are currently developing an iterative approach based on the definition of a Jacobian matrix that relates the change in MFs value with the modification applied with WUs. We rely on previous works on inverse kinematic (IK), like (Aristidou et al. 2018) that summarize different numerical approaches used to find solution on the IK problem based on the formulation of a Jacobian matrix

(inverse, pseudo-inverse, damped least squares, etc ...) The entries of our matrix are partial derivative of motions features function over a single warping unit control value (*k*) $\frac{\delta MF}{\delta k_i}$.

4.3.4 Expressivity:

Until now our analysis focuses only on the study of the pure motion, not considering any expressive interpretation of it. The final step will be computing the link between a subset of the defined MFs and a desired expressive cue. In this way we will be able to give a higher level of control over the expressivity of the motion. We are planning on doing this by performing a participants study, or by comparing real and generated motions.

4.3.5 Pynimation:

The filter is currently developed alongside with *Pynimation*, a open-source python tool for character animation developed internally. *Pynimation* is conceived as a modular software where the main functionality is: reading animation files (for now fbx, bvh and mvnx), generating an accessible and readable skeleton structure, and finally exporting again to file. On top of the basic structure we are developing different modules, exposing different functionalities like visualizing (we have developed a pyopengl module, supporting also meshes), cleaning-up motion data (like footskating), motion data analysis (like for computing motion similarities), motion editing and more. The architecture of *Pynimation* is detailed in Figure 14.
The *Expressive animation filter* itself is developed as a *Pynimation* module, the main advantage of using *Pynimation* is the abstraction from the input animation data, relying only on an abstract representation of the skeletal animation, and so the definition of a environment that can be easily integrated and implemented in different software (like Unreal and Unity).

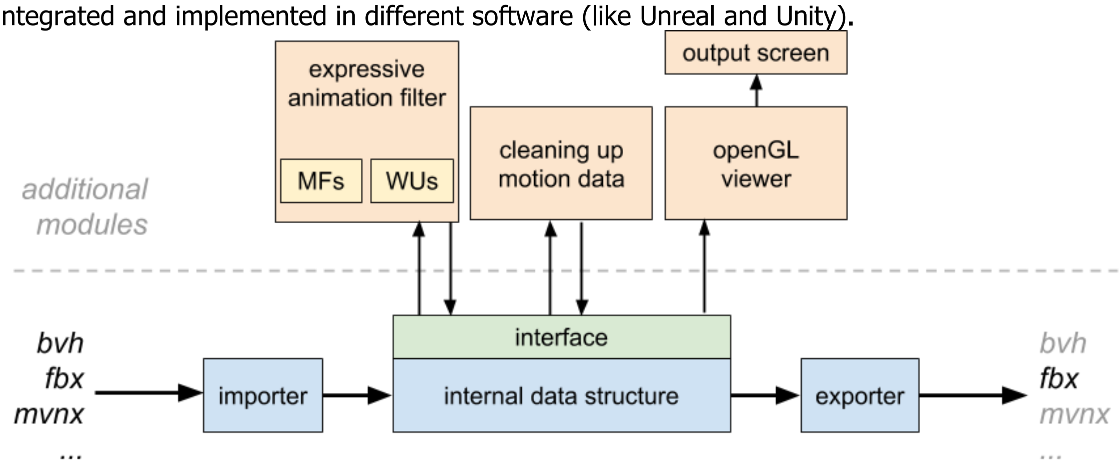


Figure 14 - Current structure of *Pynimation* with the basic structure on the bottom, and the some of the current available and developed additional modules on top (including the expressive animation filter).

4.4 Haptic techniques

In this section, we describe how we put in practice the haptic rendering techniques we have set as part of immersive technologies as shown in section 4.1. More specifically, we explore their possible effect on users behaviours in 1-to-n scenarios

4.4.1 Preliminaries

In order to explore multimodal 1-to-n interactions, we also investigated the use of wearable haptics as a modality to communicate novel types of information from the virtual agents to the user. Wearable haptic interfaces are an easy and unobtrusive way to convey contact sensations, as they can provide rich information without impairing the user's motion. As haptics is a prominent sense in our lives, we expect its use to be paramount to convey realistic and compelling interactions. As previous works have mostly been limited to distant interactions between groups of characters, due to the difficulty of rendering realistic sensations of collisions in VR, our first goal was to evaluate whether rendering physical contacts through the use of wearable haptics could influence user behaviours.

To this aim, we conducted a VR experiment (see Figure 15) where participants navigated in a crowded virtual train station, while being equipped with a motion capture suit (to capture and display their movements on their avatar), wearable haptics armbands (to render physical contacts), and a wide-field-of-view head-mounted display (HMD). Participants either experienced haptic feedback when they collided with virtual characters (i.e., when they virtually entered in contact with them), or did not receive any feedback. The purpose of the study was to investigate the effect of haptic rendering of collisions on participants' behaviour during navigation through a static crowd in VR. To explore this question, we immersed participants in a virtual trainstation and asked them to perform a navigation task which involved moving through a crowd of virtual characters. In some conditions, collisions with the virtual characters were rendered to participants using 4 wearable vibrotactile haptic devices (actuated armbands). Our general hypothesis is that haptic rendering changes the participants' behaviour by giving them feedback about the virtual collisions. Moreover,we also expect that even after removing haptic rendering, an after-effect still persists on the participants' behaviour.
Our results show that providing haptic feedback improved the overall realism of the interaction between the user and the virtual characters. In particular, participants more actively avoided collisions with the virtual characters when they experienced haptic rendering of contacts, therefore demonstrating changes in their localized interactions but no changes in their global trajectories. We also noticed a significant after-effect in the user's behavior, where they continued to show more careful interactions with the virtual characters even though they were not experiencing haptic rendering of collision in the last part of the experiment. This experiment was recently submitted to the journal IEEE Transactions on Visualization and Computer Graphics, a renowned high-impact publication in the field. These results confirmed that haptic rendering can therefore be relevant to communicate novel types of information from the virtual agents to the user, which we want to further explore to create more expressive virtual characters.

Figure 15 - We investigated the influence of rendering physical contacts while navigating in a virtual crowd. Participants wore a Xsens motion capture suit, wearable haptics armbands, and a wide-field-of-view HMD.

4.4.2 Materials & Methods

4.4.2.1 Apparatus

For the purpose of immersing participants in the virtual environment and investigating the potential effects of haptic rendering while navigating in groups of characters, we used the setup described in Section 4.1, that we briefly remind here, with more

- Motion Capture: to record participants' body motions, as well as to render their animated avatar in the scene, we used an IMU-based (Inertial Measurement Unit) motion capture system (Xsens 2).
- HMD: to immerse participants in the virtual environment, we chose to use a Pimax 3 virtual reality headset, in particular because of the wide field of view provided in these situations of close proximity with other characters
- Haptic Rendering: to render haptic collisions between participants and the virtual characters, we equipped participants with four armbands (one on each arm and forearm). Each armband is composed of four vibrotactile motors with vibration frequency range between 80 and 280 Hz and controlled independently. Motors are positioned evenly onto an elastic fabric strap. An electronics board controls the hardware. It comprises a 3.3 V Arduino Mini Pro, a 3.7 V Li-on battery, and a Bluetooth 2.1 antenna for wireless communication with the external control station.

4.4.2.2 Environment and Task

Participants were immersed in a digital reproduction of the metro station "Mayakovskaya" in Moscow, amongst a virtual static crowd (see Figure 15). A total of 8 different configurations of the scene were prepared in advance and used in the experiment. A configuration is defined by the exact position of each crowd character in the virtual station. In each configuration, the crowd formed a squared shape, and character positions followed a Poisson distribution resulting in a density of 1.47 ± 0.06 character/m^2 . Such a distribution combined with such a level of density ensures that a gap of 0.60 m on average exists between each character. The crowd is composed of standing virtual characters animated with various idle animations (only small movement but standing in place). In each configuration, characters were animated according to two types of behaviour, either waiting (oriented to face the board displaying train schedules, moving slightly the upper body) or phone-calling (with a random orientation). We used several animation clips for each of the two behaviours, in order to prevent the exact same animation clip to be used for two different virtual characters. At the beginning of each trial, participants were initially standing at one corner of the square crowd, embodied in a gender-matched avatar. They were instructed to traverse the crowd so as to reach the board displaying train schedules, and to read aloud the track number of the next train displayed on the board before coming back to their initial position. They were physically walking in the real room, while their position and movements were used to animate their avatar. This task required participants to reach the opposite corner of the space in order to read information on the board, while forcing them to move through the virtual crowd. Also, the screen displayed the train information only when participants were at less than 2 m from it (i.e., when they reached the green area displayed in Figure 5.b). Furthermore, we provided the following instruction to participants prior to the experiment: "*Walk through the virtual train station as if you were walking in a real train station*".

4.4.2.3 Protocol and Participants

Participants were equipped with the Xsens suit, the four armbands for haptic rendering, a wearable backpack computer, the head-mounted display and headphones for sound immersion. Calibration of the Xsens motion capture system was then performed to ensure motion capture quality, as well as to resize the avatar to participants dimensions. Once ready, participants performed a training trial in which they could explore the virtual environment and get familiar with the task. The experiment then consisted of 3 blocks of 8 trials, where the blocks were presented for all participants in the following order: NoHaptic1, Haptic, and NoHaptic2. The Haptic block corresponded to performing the task with haptic rendering of contacts, while the NoHaptic blocks did not involve any haptic rendering of contacts. The experiment therefore consisted in performing first a block without haptic rendering, in order to measure a baseline of participants' reactions. The purpose of the second block was then to investigate whether introducing haptic

rendering influenced their behavior while navigating in a crowd, while the purpose of the last block (without haptic) was to measure potential after-effects.
Twenty-three unpaid participants, recruited via internal mailing lists amongst students and staff, volunteered for the experiment.

4.4.2.4 Hypotheses

H1: Haptic rendering will not change the path followed by participants through the crowd. Indeed, pedestrians mainly rely on vision to control their locomotion, and we replicated each crowd configuration across the 3 blocks, resulting in identical visual information for participants to navigate. Therefore the followed path will be similar in the tree blocks of the experiment (NoHaptic1, Haptic and NoHaptic2).

H2: Haptic rendering of collisions will make participants aware of collisions and influence their body motion during the navigation through the crowd. Therefore, concerning the NoHaptic1 and Haptic blocks of the experiment, we expect that:
H2_1: Participants will navigate in the crowd more carefully in the Haptic block in order to avoid collisions. There will be more local avoidance movements (e.g., increased shoulder rotations) and a difference in participants' speed.
H2_2: With these changes on participants' local body motions, there will be both less collisions, and smaller volumes of interpenetration when a collision occurs.

H3: We expect some after-effect due to haptic rendering, i.e., we expect that participants will remain more aware and careful about collisions even after we disabled haptic rendering. Therefore we expect H2_1 and H2_2 to remain true in the NoHaptic2 block.

H4: Haptic rendering will improve the sense of presence and the sense of embodiment of participants in VR, as they will become more aware of their virtual body dimensions in space with respect to neighbour virtual characters.

4.4.3 Results

This section presents the results of our experiment, starting with the study of H1 on the trajectories formed by participants through the virtual crowd. We then explore H2_1 and H2_2 with respect to the analysis of body movements. Finally, we report the results on collision metrics so as to evaluate H3, to finish with the answers to the Presence and Embodiment questionnaires related to H4.

4.4.3.1 Trajectory Analysis

To study H1, we compared participants' trajectories through the virtual crowd. To this end, we decomposed the environment into cells based on a Delaunay triangulation, the vertices of which were the crowd characters. A trajectory is then represented as a sequence of traversed cells.

Table 1 shows the results of the Dice similarity measure between all possible pairs of blocks. Similarity ranges from 84.7% (Nohaptic1 vs. Haptic blocks) to 88.5% (Haptic vs. NoHaptic2 blocks). The score is higher for Haptic vs. Nohaptic2 blocks (88.6 ± 4.1%) and for Nohaptic1 vs. Nohaptic2 (85.9 ± 4.0%). Because it is difficult to identify from this data only whether the obtained level of similarity is due to natural variety in human behaviours, or to the difference in conditions explored in each block, we propose to measure similarity between paths belonging to the same block as follows. For each block and each configuration, we randomly divided the trajectories into two subsets and computed the Dice similarity score between them. We repeated this process 30 times (which changes the way trajectories are divided into 2 subsets). Performing this process and computing similarity over the 3 blocks resulted into 90 measures of "intra-block similarity". The obtained average value is 81.2 ± 3.3%, that can be compared with the "inter-block similarity" scores presented in Table 1. Our results show that there is no statistical difference between intra-block and Nohaptic1 vs. Haptic blocks similarity measure ($p > 0.05$. There is however a significant difference between intra-block and Haptic vs. Nohaptic2 blocks ($p < 0.01$), as well as intra-block and Nohaptic1 vs. Nohaptic2 ($p < 0.05$), where intra-block similarity measures are always lower. Given that similarity measures between pairs of blocks were either as similar or more similar than intra-block similarities, we can conclude that participants chose their path through the crowd similarly, irrespective of the block condition, which supports H1.

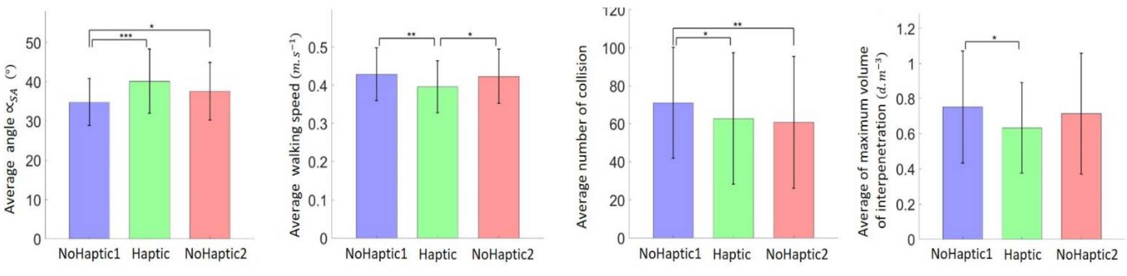
Trials	Blocks		
	NoHaptic1 vs. *Haptic*	*Haptic* vs. *NoHaptic2*	*NoHaptic1* vs. *NoHaptic2*
T_1	84.0%	88.6%	85.0%
T_2	88.4%	93.8%	88.3%
T_3	78.1%	93.2%	79.4%
T_4	91.9%	88.7%	90.7%
T_5	88.4%	90.2%	85.3%
T_6	82.8%	85.8%	91.0%
T_7	78.8%	81.6%	82.0%
T_8	85.0%	85.9%	85.3%
$\boldsymbol{T_{all}}$	**84.7 ± 4.8%**	**88.6 ± 4.1%**	**85.9 ± 4.0%**

Table 1 - Similarity measure (Dice) of participant trajectories between all blocks (NoHaptic1, Haptic, NoHaptic2) for all the trials.

4.4.3.2 Body Motion

Shoulder Rotation (Fig. 16.a). The average amplitude of shoulder rotations was significantly different in each block ($F(2, 44) = 13.0$, $p < 0.001$). In particular, it was significantly higher in the block with haptic rendering (40.1 ± 8.2 deg), than in the first block without haptic rendering (34.3 ± 6.0 deg). A higher shoulder rotation angle means that participants made a larger rotation to squeeze between virtual characters, therefore validating the hypotheses H2_1. Furthermore, it was also significantly higher in block NoHaptic2 (38.7 ± 3.7 deg) than in block NoHaptic1, suggesting that participants continued to turn more their shoulders even after haptic rendering was disabled, therefore supporting H3.

Walking Speed (Fig. 16.b). We found an effect of haptic rendering ($F(1.56, 34.2) = 7.14$, $p = 0.005$) on participant's average walking speed, where participants' walking speed was on average significantly lower in theHaptic block (0.40±0.07 m/s) than in the NoHaptic1 (0.43±0.07 m/s) and NoHaptic2 (0.42±0.07 m/s) blocks. This result therefore supports hypothesis H2_1 .



a) shoulder rotation
b) walking speed
c) number of collisions
d) volume of interpenetrati on

Figure 16 – Analysis results for collected data

4.4.3.3 Number of collisions are volume of interpenetration

We analyzed the number of collisions as well as the volume of interpenetration between the user's avatar and the virtual agents in the scene, shown in Figs. 16.c and 16.d. The average number of collisions per trial was influenced by haptic rendering with a large effect ($F(2, 44) = 7.13$, $p = 0.002$). Post-hoc analysis showed that the number of collisions was higher during the NoHaptic1 block (71 ± 29.2) than during the Haptic (62.8 ± 34.6 , $p = 0.018$) and NoHaptic2 blocks (60.7 ± 34.6, $p = 0.002$), which shows that participants made on average more collisions before they experienced haptic rendering. The average volume of interpenetration was also influenced by the block ($F(2, 44) = 4.35$, $p = 0.019$), where post-hoc analysis showed that this volume was smaller ($p = 0.016$) in the Haptic block (0.6 ± 0.3 dm^{-3}) than during the NoHaptic1 (0.8 ± 0.3 dm^{-3}). These results validate our hypothesis H2_2, that states that haptic rendering reduces the severity of collisions between participants and virtual characters. Furthermore, as the number of collisions is higher during block NoHaptic1 than during block NoHaptic2, this also supports H3 on potential after-effects of haptic rendering.

4.4.3.4 Presence and Embodiment

Presence and Embodiment. Another important aspect of our analysis is its perceptual relevance. In accordance with H4, we looked for any difference in the users' feelings of presence and embodiment, comparing the registered subjective perception with and without haptic rendering. Participants answered both questionnaires at the end of each block (Embodiment then Presence), answering each question on a 7-point Likert scale.

The average participant ratings and all the questions for embodiment are shown in Tables 2, 3, and 4. We did not find any significant effect of the blocks for Agency ($p = 0.438$), Change ($p = 0.085$) and Ownership ($p = 0.753$). Furthermore, Table 5 shows the questions and the average participant ratings for presence, for which we also did not find a significant effect of the blocks ($p = 0.222$). These results therefore do not support hypothesis H4, suggesting that haptic rendering does not improve the sense of presence or the sense of embodiment of participants in VR.

Questions	**blocks**		
	NoHaptic1	*Haptic*	*NoHaptic2*
The movements of the virtual body felt like they were my movements. I felt like I was controlling the movements of the virtual body I felt like I was causing the movements of the virtual body. The movements of the virtual body were in sync with my own movements.	6.1 ± 0.9	6.0 ± 0.8	5.9 ± 0.7

Table 2 - Agency questionnaire: average participant ratings for the three blocks.

Questions	Blocks		
	NoHaptic1	*Haptic*	*NoHaptic2*
I felt like the form or appearance of my own body had changed. It felt like the weight of my own body had changed. I felt like the size (height) of my own body had changed. I felt like the width of my own body had changed.	3.6 ± 1.3	3.8 ± 1.5	3.3 ± 1.5

Table 3 - Ownership questionnaire: average participant ratings for the three blocks.able 3-

Questions	Blocks		
	NoHaptic1	*Haptic*	*NoHaptic2*
It felt like the virtual body was my body. It felt like the virtual body parts were my body parts. The virtual body felt like a human body. It felt like the virtual body belonged to me.	4.9 ± 1.4	5.1 ± 1.2	5.0 ± 1.2

Table 4 - Slater-Usoh-Steed (SUS) questionnaire and average participant ratings for the three blocks.

Questions	Blocks		
	NoHaptic1	*Haptic*	*NoHaptic2*
I had a sense of being there in the train station. There were times during the experience when... the train station was the reality for me... The train station seems to me to be more like... I had a stronger sense of... I think of the train station as a place in a way similar to other places that I've been today. During the experience I often thought that I was really standing in the train station.	5.2 ± 0.9	5.2 ± 1.2	5.0 ± 1.1

Table 5 - Change questionnaire: average participant ratings for the three blocks.

4.4.4 Discussion

The main objective of this study was to evaluate the effect of haptic rendering of collisions on participants' behaviour while navigating in a dense virtual crowd. These results confirmed that haptic rendering can therefore be relevant to communicate novel types of information from the virtual agents to the user, which we want to further explore to create more expressive virtual characters.

Trajectories. The analysis of the Dice similarity measure showed that haptic rendering did not change the way participants selected their path through the crowd, as stated in hypothesis H1 . We even found that paths across blocks were "more similar" than within the same block. One possible explanation is given by the way we compose the sets we compare the similarity of, where we assume that paths are independent from participants. Indeed, the intra-block similarity measure required us to split a set of trajectories belonging to the same block and crowd configuration, which resulted into comparing paths performed by different participants. In contrast, the inter-block analysis considered sets that were split according to haptic rendering conditions, thus comparing paths performed by the same group of 23 participants. In spite of this limitation in our analysis, we consider that paths are similar across blocks. One can describe human motion as a trajectory resulting from a perception-action loop. Depending on the tasks, the loop is a multi-modal one, meaning that different senses are used to control motion. However in the context of walking, vision is the most used perceptual input to navigate to the goal. Such statements hold in our case, where a major difference with previous work is the higher density of obstacles. Nevertheless, assuming that tactile feedback may affect path selection, it would have been probable that some participants reversed their course after a collision has been rendered, which was not observed.

Avoidance Behaviour. In this experiment, we demonstrated that haptic rendering had an effect on shoulder rotations, which supports hypothesis H2_1. In particular, participants rotated more their shoulders when traversing the gaps between virtual characters during the Haptic block than during the NoHaptic1 block. Let us remind that the human trunk is most often larger along the transverse axis than along the antero-posterior axis. Thus, the more the participants turn their shoulders the smaller the volume swept by their body motion. Our results therefore suggest that participants might have tried to minimize the risk of collision with virtual characters more in the condition where they experienced haptic rendering than in the first block of the experiment. The slower speed observed in the Haptic block also reveals that participants moved more cautiously. Being more cautious effectively resulted in less collisions as expected in hypothesis H2_2. Results show that the average number of collisions as well as the average volume of interpenetration were significantly lower in the Haptic block than in the NoHaptic1 block.

Haptic Rendering After-effects. While there were less collisions and more shoulder rotations observed in the Haptic block in comparison with the NoHaptic1 block, there was no difference between the Haptic and the NoHaptic2 blocks. This supports hypothesis H3 on potential after-effects of haptic rendering. However, such an after-effect did not equally influence all measurements, such as walking speed that increased again in the NoHaptic2 block. One possible explanation might be a perceptual calibration of the participants. During the experiment, participants became more familiar with the environment, the task to be performed, but also the virtual representation of their body and the virtual environment, enabling them to move faster and better avoid collisions with the virtual characters in the last block (NoHaptic2). Another point to highlight is that participants, at the beginning of the Haptic block, did not know that contacts would now trigger a vibrotactile haptic sensation. For this reason, we might expect to see a short learning phase at the beginning of the block, where participants learn to deal with the newly-rendered haptic collisions. Considering this point, we can expect the effect of providing haptic sensations of collisions even stronger than registered. However, to provide a more definitive conclusion on the role of the haptic after-effect would require to add a control group with no haptic rendering throughout the 3 blocks of the experiment, which could be explored in future work. These results can also open perspectives regarding the design of new experiments including haptic priming tasks.

Embodiment and Presence. In contrast with our hypothesis H4, we did not find any significant change in terms of the user's perceived senses of embodiment and presence when experiencing haptic feedback. This result is quite surprising, as we did find significant effects in other measurements, suggesting that participants took different actions when provided with haptic sensations of contact. An explanation for this result could lie in the fact that users already registered high embodiment and presence levels without experiencing haptic feedback in the first condition (NoHaptic1), leaving little room for improvement in the Haptic condition. Finally, a last explanation could be the location and number of our haptic devices. Employing a higher number of bracelets spread throughout the body might better render the target contact sensations. All these considerations will drive our future work.

4.4.5 Summary and future works

In this work, we designed an experiment to evaluate the effects, as well as the after-effects, of haptic rendering on a motion task in a highly crowded environment. Participants performed a goal-directed navigation task through a dense virtual crowd. Wearable haptic devices provided them with vibrotactile feedback whenever a collision with their arms occurred. Results showed that providing haptic feedback impacted the way participants moved through the virtual crowd. They were more cautious about the collisions they provoked with virtual characters, but they did not change their global trajectories. We also demonstrated the presence of an after-effect of haptic feedback, since changes in their movements remained after haptic feedback was disabled. Finally, quite surprisingly, we did not notice any impact of haptic rendering on the perceived Presence and Embodiment. These results show that visual information is probably the main sense used for navigation in dense crowds. However, a combination of visual and haptic feedback improves the overall realism of the experience, as participants show a more realistic behaviour: they are more cautious about not touching virtual characters. For this reason, we therefore suggest using haptic rendering to study human behaviour and locomotion interactions that may lead to contacts.

For future work, we are interested in populating our virtual environments with more interactive and reactive virtual characters. This is a crucial aspect since it seems to be a requirement to further improve the feeling of presence of participants. Also, the use of reactive characters may increase the effect of haptic rendering, since we could expect stronger participant reactions when virtual characters would also react after a collision. A more detailed analysis that evaluates motion before and after a collision is rendered and a virtual character reacts would then also be relevant to study. We are also interested in carrying out experiments enrolling more subjects and analysing a wider range of metrics in different scenarios (e.g., considering a dynamic crowd, measuring the effect on shoulder hunching, carrying out a control experiment where no haptics is applied). Finally, we plan to use more compelling wearable haptic devices to provide a more realistic sensation of collision while keeping the overall system compact and easy to wear.

5 CONCLUSION

In this first period of the PRESENT project, we have set strong foundations for building a solution to equip PRESENT agents with reactive and expressive capabilities through collective movements, body movement as well as tactile sensory channels.

We also conducted first evaluation of the targeted technologies, so as to confirm the relevance of the working directions we have set.

In the following period of the project, we have set new targets for our work:

- Concerning Interaction Fields, our objective is now to put this technique in the hand of designers, to evaluate, through users studies, new possibilities in elaborating complex interactive scenarios in a collective context.
- Concerning the expressive animation filter, now that we have set the technical foundations for this approach, we need to elaborate a correspondence between a set of expressions and reactions we would like to see happen in interaction scenarios, with a solid list of motion features we can control efficiently. The efficacy of the completed animation system will be then evaluated through perceptual studies in an immersive context.

6 REFERENCES

Argyle, M. (1969). Social interaction. Tavistock publications.

Aristidou, A., Lasenby, J., Chrysanthou, Y., & Shamir, A. (2018, September). Inverse kinematics techniques in computer graphics: A survey. In Computer Graphics Forum (Vol. 37, No. 6, pp. 35-58).

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. Organizational influence processes, 295-303.

Bailenson, J.N., Blascovich, J., Beall, A. C., and Loomis, J. M. (2003). Interpersonal distance in immersive virtual environments. Personality and Social Psychology Bulletin, 29, 7, 819–833.

Brady, A. and Walker, M.B. (1978). Interpersonal distance as a function of situationally induced anxiety. British Journal of Social and Clinical Psychology, 17, 2, 127–133.

Cafaro, A., Ravenet, B., Ochs, M., Vilhjálmsson, H. H., & Pelachaud, C. (2016). The effects of interpersonal attitude of a group of agents on user's presence and proxemics behavior. ACM Transactions on Interactive Intelligent Systems (TiiS), 6(2), 1-33.

Christian, J.J., Vagh F., and Davis, D. E. (1961). Phenomena associated with population density. Proceedings of the National Academy of Science 47:428-49.

Dael, N., Mortillaro, M., & Scherer, K. R. (2012). The body action and posture coding system (BAP): Development and reliability. Journal of Nonverbal Behavior, 36(2), 97-121.

Ekman, P., & Friesen, W. V. (1978). Facial action coding systems. Consulting Psychologists Press.

Hall, E. T. (1963). A system for the notation of proxemic behavior. American anthropologist, 65(5), 1003-1026.

Hall, E. T., Birdwhistell, R. L., Bock, B., Bohannan, P., Diebold Jr, A. R., Durbin, M., and La Barre, W. (1968). Proxemics [and comments and replies]. Current anthropology, 9(2/3), 83-108.

Helbing, D., & Molnar, P. (1995). Social force model for pedestrian dynamics. Physical review E, 51(5), 4282.

Iachini, T., Coello, Y., Frassinetti, F., and Ruggiero, R. (2014). Body space in social interactions: a comparison of reaching and comfort distance in immersive virtual reality. PloS one, 9, 11, e111511.

Juslin, P. N., Scherer, K. R., Harrigan, J. A., & Rosenthal, R. (2005). The new handbook of methods in nonverbal behavior research.

Kendon, A. (1990). Conducting interaction: Patterns of behavior in focused encounters (Vol. 7). CUP Archive.

Kmiecik, C., Mausar, P., & Banziger, G. (1979). Attractiveness and interpersonal space. The Journal of Social Psychology, 108(2), 277-278.

Remland, M. S., Jones, T. S., & Brinkman, H. (1995). Interpersonal distance, body orientation, and touch: Effects of culture, gender, and age. The Journal of social psychology, 135(3), 281-297.

López, A., Chaumette, F., Marchand, E., & Pettré, J. (2019, May). Character navigation in dynamic environments based on optical flow. In Computer Graphics Forum (Vol. 38, No. 2, pp. 181-192).

Patil, S., Van Den Berg, J., Curtis, S., Lin, M. C., & Manocha, D. (2010). Directing crowd simulations using navigation fields. *IEEE transactions on visualization and computer graphics*, *17*(2), 244-254.

Randhavane, T., Bhattacharya, U., Kapsaskis, K., Gray, K., Bera, A., & Manocha, D. (2019). Identifying emotions from walking using affective and deep features. arXiv preprint arXiv:1906.11884.

Santello, M., Bianchi, M., Gabiccini, M., Ricciardi, E., Salvietti, G., Prattichizzo, D., ... & Kyriakopoulos, K. (2016). Hand synergies: Integration of robotics and neuroscience for understanding the control of biological and artificial hands. Physics of life reviews, 17, 1-23.

Slater, Mel & Pertaub, David-Paul & Barker, Chris & Clark, David. (2006). An Experimental Study on Fear of Public Speaking Using a Virtual Environment. Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society. 9. 627-33. 10.1089/cpb.2006.9.627.

Sweeny, T. D., & Whitney, D. (2014). Perceiving Crowd Attention: Ensemble Perception of a Crowd's Gaze. Psychological Science, 25(10), 1903–1913. https://doi.org/10.1177/0956797614544510

Treuille, A., Cooper, S. & Popović, Z. (2006). Continuum crowds. <i>ACM Trans. Graph.</i> 25, 3 (July 2006), 1160–1168.

van den Berg, Jur & Guy, Stephen & Lin, Ming & Manocha, Dinesh. (2011). Reciprocal n-Body Collision Avoidance. 10.1007/978-3-642-19457-3_1.

Wilder, D. A. (1986). Social categorization: Implications for creation and reduction of intergroup bias. In Advances in experimental social psychology (Vol. 19, pp. 291-355). Academic Press.

Witkin, A., & Popovic, Z. (1995, September). Motion warping. In Proceedings of the 22nd annual conference on Computer graphics and interactive techniques (pp. 105-108).